

*Citation for published version:*

Gibbs, T, Grilli, J, Rogers, T & Allesina, S 2018, 'Effect of population abundances on the stability of large random ecosystems', *Physical Review E*, vol. 98, no. 2, 022410, pp. 1 - 16. <https://doi.org/10.1103/PhysRevE.98.022410>

*DOI:*

[10.1103/PhysRevE.98.022410](https://doi.org/10.1103/PhysRevE.98.022410)

*Publication date:*

2018

*Document Version*

Peer reviewed version

[Link to publication](https://doi.org/10.1103/PhysRevE.98.022410)

Copyright ©2018 American Physical Society. The final publication is available at PHYSICAL REVIEW E via <https://doi.org/10.1103/PhysRevE.98.022410>

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# The effect of population abundances on the stability of large random ecosystems

Theo Gibbs,<sup>1</sup> Jacopo Grilli,<sup>1</sup> Tim Rogers,<sup>2</sup> and Stefano Allesina<sup>1,3,4</sup>

<sup>1</sup>*Dept. of Ecology & Evolution, University of Chicago. Chicago, IL 60637, USA*

<sup>2</sup>*Centre for Networks and Collective Behaviour, Department of Mathematical Sciences,  
University of Bath. Claverton Down Bath BA2 7AY United Kingdom.*

<sup>3</sup>*Computation Institute, University of Chicago.*

<sup>4</sup>*Northwestern Institute on Complex Systems (NICO), Northwestern University.*

Random matrix theory successfully connects the structure of interactions of large ecological communities to their ability to respond to perturbations. One of the most debated aspects of this approach is that so far studies have neglected the role of population abundances on stability. While species abundances are well-studied and empirically accessible, studies on stability have so far failed to incorporate this information. Here, we tackle this question by explicitly including population abundances in a random matrix framework. We derive an analytical formula that describes the spectrum of a large community matrix for arbitrary feasible species abundance distributions. The emerging picture is remarkably simple: while population abundances affect the rate to return to equilibrium after a perturbation, the stability of large ecosystems is uniquely determined by the interaction matrix. We confirm this result by showing that the likelihood of having a feasible and unstable solution in the Lotka-Volterra system of equations decreases exponentially with the number of species for stable interaction matrices.

## I. INTRODUCTION

Since the work of Lotka and Volterra, ecologists have attempted to mathematize the interactions between populations to build predictive models of population dynamics. This is a complex problem: the equations describing their interactions have been debated for decades [1], ecological communities are often composed of a large number of species [2], and the estimation of parameters and initial conditions is often unfeasible.

To circumvent this problem, Robert May [3] introduced the idea of modeling complex ecological communities using random matrices. Consider the case in which the dynamics of the populations can be described by a system of ordinary differential equations:

$$\frac{dx_i(t)}{dt} = f_i(\mathbf{x}(t)), \quad (1)$$

where  $\mathbf{x}(t)$  is a vector containing the populations abundances at time  $t$ , and the function  $f_i$  relates the abundance of all populations to the growth of population  $i$ . In general,  $f_i$  is a nonlinear equation with several parameters.

Suppose that the system admits a feasible equilibrium point, i.e., a vector  $\mathbf{x}^*$  such that  $f_i(\mathbf{x}^*) = 0$  and  $x_i^* > 0$  for all  $i$ . If we start the system at this point, it will remain there indefinitely. We can therefore ask whether the system will go back to the equilibrium, or rather move away from it, following a perturbation. This type of stability analysis can be carried out by building the Jacobian matrix  $J_{ij} = \partial f_i(\mathbf{x}(t))/\partial x_j$  and evaluating it at the equilibrium point, yielding the so-called community matrix  $\mathbf{M} = \mathbf{J}|_{\mathbf{x}^*}$ . If all the eigenvalues of  $\mathbf{M}$  have negative real part, then the equilibrium is locally asymptotically stable, and the system will return to it after sufficiently small perturbations; if any of the eigenvalues have a positive real part, the system will move away from the equilibrium when perturbed.

Clearly, to build  $\mathbf{M}$  one would need to precisely know the functions  $f_i$ , as well as their parameters, and solve for the equilibrium (or equilibria)  $\mathbf{x}^*$ . May took a radically different approach and analyzed the case in which  $\mathbf{M}$  is a random matrix with independent, identically distributed off-diagonal elements, and constant diagonal elements [3]. For this parameterization, he was able to show that the community matrices describing sufficiently large and complex ecological communities are always unstable. The random-matrix approach was recently extended and refined to include different types of interaction between the populations [4, 5], as well as to study the effect of more complex network structures, such as the hierarchical organization of food-webs [6] and the modular pattern often displayed by biological networks [7].

By modeling directly the matrix  $\mathbf{M}$  as a random matrix, one does not require a precise characterization of the functions  $f_i$  and the equilibrium  $\mathbf{x}^*$ . While mathematically convenient, this approach does not explicitly take into account the abundance of the populations—a quantity that is empirically much more accessible than interaction coefficients or the elements of the community matrix.

The distribution of species abundances (SAD) is one of the most studied patterns in ecology and it has been shown to have remarkably similar features across different species-rich communities [8] with a skewed shape and few highly abundant species. The log-series distribution [9], discrete lognormal [10] and negative binomial [11] have all been proposed to describe empirical SADs, and have been shown to emerge from either neutral [12–15] or niche mechanisms [16, 17]. An important limitation of these studies is that typically species abundances are studied and modeled by considering a single trophic level or a single taxonomic group.

The role of species abundances in structuring the community matrix  $\mathbf{M}$  can be easily seen when considering one of the simplest models of population dynamics, the Generalized Lotka-Volterra (GLV) model:

$$\frac{dx_i(t)}{dt} = x_i(t) \left( r_i + \sum_j A_{ij} x_j(t) \right), \quad (2)$$

where  $r_i$  is the intrinsic growth rate of species  $i$ , and  $A_{ij}$  is the per-capita effect of species  $j$  on the growth of  $i$ . If a feasible equilibrium (i.e., one where all species have positive abundance) exists, then it can be found solving the system of equations

$$0 = r_i + \sum_j A_{ij} x_j^*, \quad (3)$$

yielding the community matrix  $M_{ij} = A_{ij} x_i^*$ , which can be written in matrix form as

$$\mathbf{M} = \mathbf{X} \mathbf{A}, \quad (4)$$

where  $\mathbf{X}$  is a diagonal matrix with  $X_{ii} = x_i^*$  and zeros elsewhere. Even if the elements of  $\mathbf{A}$  were independent, identically distributed samples from a distribution, the elements of  $\mathbf{M}$  would not be—the matrix of abundances  $\mathbf{X}$  couples all the coefficients in the same row, such that the distribution of the elements in each row would in principle be different.

One of the main goals of this work is to extend the random matrix approach by considering a random matrix of abundances  $\mathbf{X}$  and a random matrix of interactions  $\mathbf{A}$ , and determining the stability of  $\mathbf{M}$  under these conditions. In this way, we address the effect of species abundances on stability, thereby lifting one of the main criticisms of the random matrix approach [5, 18, 19].

Our results extend to more complex models. For example, consider

$$\frac{dx_i(t)}{dt} = \phi_i(x_i(t)) H_i \left( \sum_j A_{ij} x_j(t) \right), \quad (5)$$

where  $\phi_i$  and  $H_i$  are positive and monotonically increasing functions. It is easy to observe that, assuming  $\phi_i(0) = 0$ , the Jacobian evaluated at the fixed point has the same form of equation 4, with  $X_{ii} = \phi_i(x_i^*) H_i' \left( \sum_j A_{ij} x_j^* \right)$ . While for simplicity we will focus on the GLV case, our approach yields insights to more complicated dynamics, and goes beyond the ecological application that motivates this study.

As we stated above, when analyzing coexistence, we need population abundances to be positive (*feasible*). Stability cannot, at least in principle, be disentangled from the constraint imposed by feasibility on interactions [20]. Diversity and interaction properties have important consequences for the range of parameters corresponding to feasible solutions [21–23]. While the interest in feasibility has grown considerably in recent years, the relationship between feasibility and stability is still unclear. In fact, most of the studies on feasibility assume strong conditions on the interaction matrix (e.g., D-stability, diagonal stability) that guarantee stability of feasible solution [21, 23] (see Supplementary Information). It is still unclear whether and when these assumptions are justified, and how likely it is for large random interaction matrices to meet these conditions.

In the second part of this work, we focus on the relationship between feasibility and stability. In particular, we study the relationship between the stability of  $\mathbf{A}$  and that of  $\mathbf{M}$  for the GLV model. Our results show that, given a stable random matrix  $\mathbf{A}$ , the probability that an arbitrary feasible equilibrium is unstable decreases exponentially with diversity. This result strongly suggests that, provided that the interaction matrix  $\mathbf{A}$  is stable, feasible solutions are almost surely stable. We therefore provide a more robust justification to both May's original paper—by showing that population abundances do not affect qualitatively stability—and the more recent work on feasibility that assumes stability—by predicting that this assumption is almost surely met for large random systems.

## II. CONSTRUCTING THE COMMUNITY MATRIX WITH ARBITRARY POPULATION ABUNDANCE

We consider a system of  $S$  interacting populations whose dynamics are described by the GLV model in equation A1, assume that a feasible equilibrium  $\mathbf{x}^*$  exists, and define  $\mathbf{X}$  as the diagonal matrix with diagonal entries  $X_{ii} = x_i^*$ . The feasible fixed point  $\mathbf{x}^*$  is locally asymptotically stable if and only if all the eigenvalues of the community matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$ , with components  $M_{ij} = x_j^* A_{ij}$ , have a negative real part. Here, we model  $\mathbf{A}$  as a random matrix and  $\mathbf{x}^*$  as a random vector with positive components, with the goal of studying the spectrum (distribution of the eigenvalues) of the community matrix  $\mathbf{M}$ . From the GLV model, specifying a feasible fixed point  $\mathbf{x}^*$  is the same as specifying a vector of intrinsic growth rates  $\mathbf{r}$  inside the feasibility domain [21, 23]. It is important to underline that the focus here is in fact different from the one of refs. [21, 23]. In particular, the goal of ref. [23] was to study the probability to find a feasible solution of a Lotka-Volterra system of equations with random interaction coefficients. Ref. [23], and many others on a similar topic, assumed a strong condition of stability on the interaction matrix. Under that assumption, any feasible solution is stable. This manuscript instead assumes that the feasibility condition holds and focuses on the stability of feasible fixed point.

More specifically, we assume that the diagonal entries of the diagonal matrix  $\mathbf{X}$  are drawn from an arbitrary distribution with positive support, mean  $\mu_X$ , and variance  $\sigma_X^2$ . The diagonal entries of  $\mathbf{A}$  are drawn from an arbitrary distribution with support in the negative axis, mean  $\mu_d$ , and variance  $\sigma_d^2$ . Finally, each off-diagonal pair  $(A_{ij}, A_{ji})$  in  $\mathbf{A}$  is drawn independently from a bivariate distribution with identical marginal means  $\mu$ , variances  $\sigma^2$ , and correlation  $\rho$ . Unless otherwise specified, we focus on the case  $\sigma_d = 0$ , while we discuss in the Supplementary Information the effects of variability in self-regulation [24].

In the case of  $\sigma_d = 0$  and in the limit of large  $S$ , the spectrum of  $\mathbf{A}$  is known and is independent of the choice of the bivariate distribution (provided that mild conditions on the finiteness of the moments are satisfied [25]). In particular,  $\mathbf{A}$  has one eigenvalue equal to  $-\mu_d + S\mu$  [26], while the others (the *bulk* of eigenvalues) are uniformly distributed in an ellipse in the complex plane centered in  $-\mu_d - \mu$  with horizontal axis  $\sqrt{S}\sigma(1 + \rho)$  and vertical axis  $\sqrt{S}\sigma(1 - \rho)$  [4, 25, 26]. Figure 1 shows an example of the spectrum of  $\mathbf{A}$ .

Figure 1 also shows an example of the eigenvalue distribution of the community matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$  where the diagonal entries of  $\mathbf{X}$  are independent random variables drawn from a uniform distribution. It is evident that the bulk of eigenvalues of  $\mathbf{M}$  does not follow the elliptic law. The main goal of this work is to characterize the spectrum of  $\mathbf{M}$  given the properties of  $\mathbf{A}$  and  $\mathbf{X}$ .

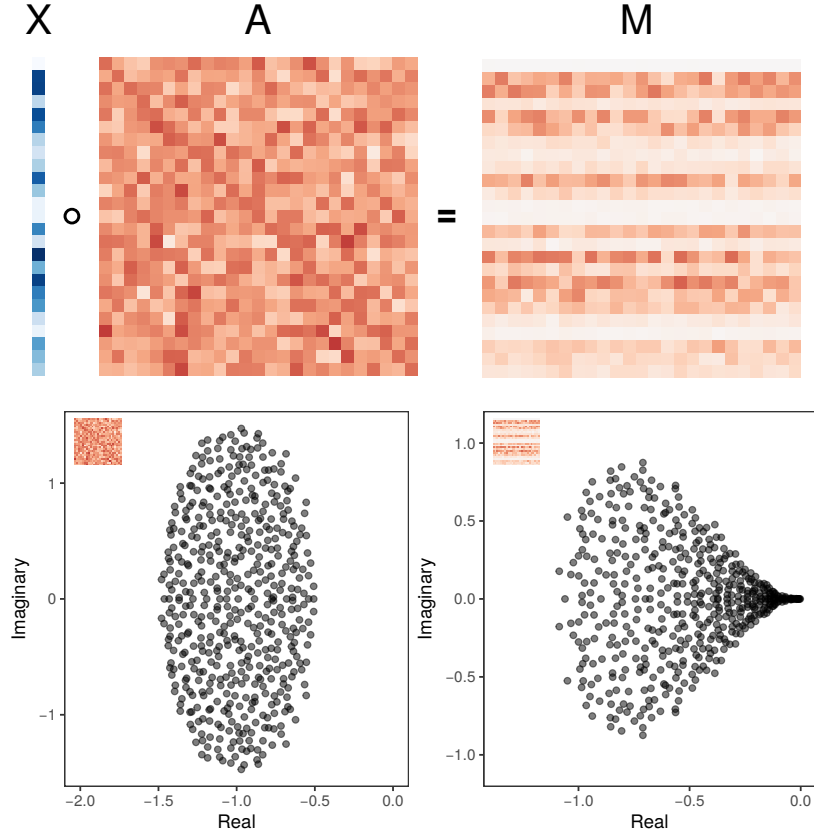


FIG. 1. The top row shows the vector of abundances  $\mathbf{x}^*$ , the interaction matrix  $\mathbf{A}$  and the community matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$  (where  $\mathbf{X}$  is a diagonal matrix with diagonal entries  $\mathbf{x}^*$ ), with colors from red (negative) to green (positive). The bottom row shows the eigenvalue distribution of  $\mathbf{A}$  and  $\mathbf{M}$ , for  $S = 500$ . The diagonal entries of  $\mathbf{X}$  are sampled from a uniform distribution on  $[0, 1]$ , and matrix  $\mathbf{A}$  is built sampling independently each pair  $(A_{ij}, A_{ji})$  from a normal bivariate distribution with identical marginals defined by  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$ , and correlation  $\rho = -0.5$ . The diagonal elements of  $\mathbf{A}$  are fixed at  $-1$ . The distribution of  $\mathbf{X}$  strongly affects the shape of the eigenvalue distribution of  $\mathbf{M}$ .

### III. DISENTANGLING THE EFFECT OF THE MEAN INTERACTION STRENGTH

When the mean  $\mu$  of the off-diagonal elements of the interaction matrix  $\mathbf{A}$  does not equal zero, the spectra of  $\mathbf{A}$  and  $\mathbf{M}$  are characterized by the presence of an outlier. The value of this eigenvalue for the matrix  $\mathbf{A}$  is known for the case  $\sigma_d = 0$ , and in the limit of large  $S$  [26]. It can be obtained by decomposing the matrix  $\mathbf{A}$  as a sum of three matrices

$$\mathbf{A} = (\mu_d - \mu)\mathbf{I} + \mu\mathbf{1} + \mathbf{B}, \quad (6)$$

where  $\mathbf{I}$  is the identity matrix,  $\mathbf{1}$  is a matrix of ones, and  $\mathbf{B}$  measures the deviation of the entries of  $\mathbf{A}$  from the mean  $\mu$ . Following our parameterization,  $\mathbf{B}$  is a random matrix with mean zero that follows the elliptic law. O'Rourke & Renfrew [26] proved that the spectrum of  $\mathbf{A}$  is characterized by a bulk of eigenvalues, determined by the spectrum of  $(\mu_d - \mu)\mathbf{I} + \mathbf{B}$ , and the presence of an outlier, whose value is (approximately) given by the largest eigenvalue of  $(\mu_d - \mu)\mathbf{I} + \mu\mathbf{1}$ , which has value  $\mu_d + (S - 1)\mu$ .

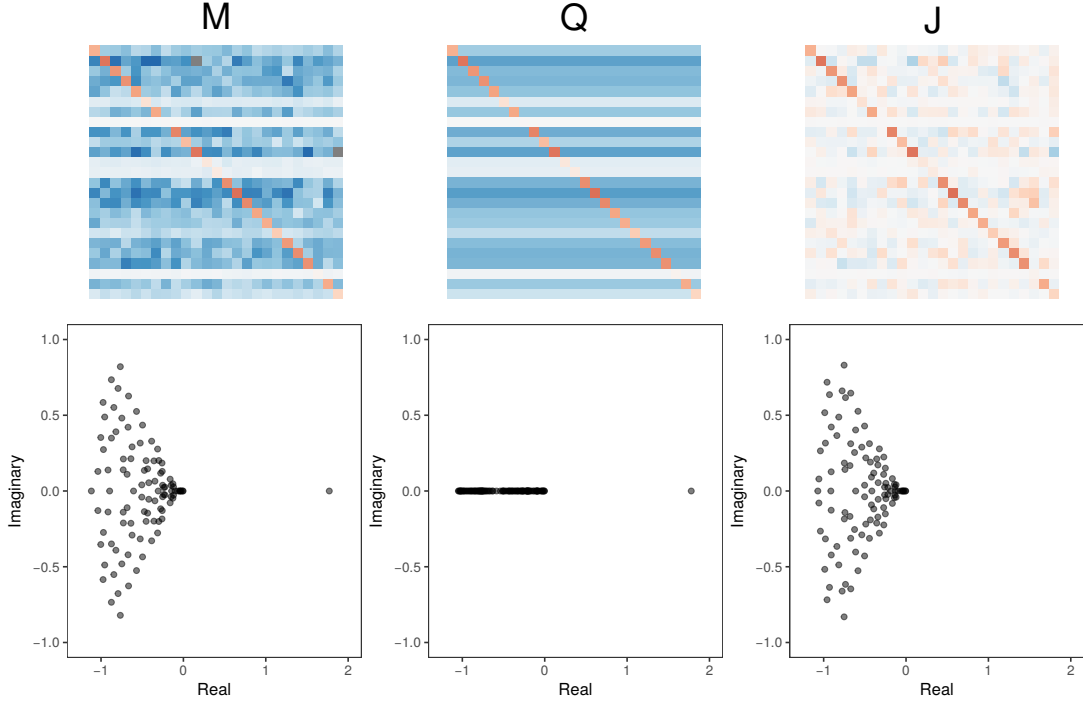


FIG. 2. The top row shows the three matrices  $\mathbf{M}$ ,  $\mathbf{Q}$  and  $\mathbf{J}$ . The community matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$ , is obtained from the interaction matrix  $\mathbf{A}$  that, without loss of generality, can be written as  $\mathbf{A} = (\mu_d - \mu)\mathbf{I} + \mu\mathbf{1} + \mathbf{B}$ , where  $\mathbf{1}$  is a matrix of ones and  $\mathbf{B}$  is a random matrix with diagonal elements fixed at zero whose coefficients have mean zero and variance  $\sigma^2$ . We define  $\mathbf{Q} = \mathbf{X}((\mu_d - \mu)\mathbf{I} + \mu\mathbf{1})$  and  $\mathbf{J} = \mathbf{X}(\mu_d\mathbf{I} + \mathbf{B})$ . Equivalently,  $\mathbf{Q}$  is the matrix with the same parameters as  $\mathbf{M}$  except with  $\sigma = 0$ , and  $\mathbf{J}$  is obtained from the same parameters as  $\mathbf{M}$  except with  $\mu = 0$  for the off-diagonal terms. The bottom rows show the eigenvalues of  $\mathbf{M}$ ,  $\mathbf{Q}$  and  $\mathbf{J}$ . Remarkably, the spectra are simply related: the bulk of the eigenvalue distributions of  $\mathbf{J}$  and that of  $\mathbf{M}$  are the same, while the outlier of  $\mathbf{M}$  is the same as that of  $\mathbf{Q}$ . In the figure, we set  $S = 500$ . The diagonal entries of  $\mathbf{X}$  are sampled from a uniform distribution on  $[0, 1]$ . The coefficients of  $\mathbf{A}$  are sampled from a normal bivariate distribution with identical marginals  $\mu = 5/S$ ,  $\sigma = 5/\sqrt{S}$  and correlation  $\rho = -0.5$ .

Figure 2 shows that, if  $\mu \neq 0$ , the spectrum of  $\mathbf{M}$  is also characterized by the presence of a bulk and of an outlying eigenvalue. By decomposing the matrix  $\mathbf{M}$  as

$$\mathbf{M} = \mathbf{X}((\mu_d - \mu)\mathbf{I} + \mu\mathbf{1} + \mathbf{B}) , \quad (7)$$

we show in the Supplementary Information that the bulk of the spectrum of  $\mathbf{M}$  is determined by the eigenvalues of the matrix  $\mathbf{J} = \mathbf{X}((\mu_d - \mu)\mathbf{I} + \mathbf{B})$  and the outlier is given by largest eigenvalue of  $\mathbf{Q} = \mathbf{X}((\mu_d - \mu)\mathbf{I} + \mu\mathbf{1})$ . Figure 2 shows an example of this decomposition, where it is evident that the bulks of eigenvalues of  $\mathbf{M}$  and  $\mathbf{J}$  are the same, and the outliers of  $\mathbf{M}$  and  $\mathbf{Q}$  match. This decomposition allows us to obtain an analytical prediction for the outlier, and in the Supplementary Information we find the spectrum of  $\mathbf{Q}$  analytically.

The trace of  $\mathbf{M}$  is given by

$$\text{tr}(\mathbf{M}) = \lambda_{\text{out}} + (S - 1)\langle \lambda \rangle_{\text{bulk}} , \quad (8)$$

where  $\lambda_{\text{out}}$  is the value of the outlier and  $\langle \lambda \rangle_{\text{bulk}}$  is the average eigenvalue in the bulk. Since the bulks of the eigenvalues of  $\mathbf{M}$  and  $\mathbf{J}$  are the same, we have that

$$\langle \lambda \rangle_{\text{bulk}} = \frac{1}{S} \text{tr}(\mathbf{J}) = \mu_X (\mu_d - \mu) . \quad (9)$$

Using the fact that

$$\text{tr}(\mathbf{M}) = S\mu_d\mu_X , \quad (10)$$

we see that the outlier is equal to

$$\lambda_{\text{out}} = \mu_X (\mu_d + (S - 1)\mu) . \quad (11)$$

Figure 3 shows that this analytical prediction closely matches the outlier of the spectrum of  $\mathbf{M}$ . We can observe deviations from our prediction when  $\mu$  is small, especially when  $\mathbf{X}$  is drawn from a log-normal distribution. This is because the eigenvalue corresponding to equation 11 is now contained in the bulk.

For mutualistic systems (where  $\mu > 0$ ), the outlier is the eigenvalue determining the stability of the matrix  $\mathbf{M}$ . Equation 11 implies that for mutualistic systems, if the matrix  $\mathbf{A}$  is stable, then the community matrix  $\mathbf{M}$  is also stable, and therefore all the feasible fixed points are expected to be stable. In this case,  $\lambda_M$ , the rightmost eigenvalue of  $\mathbf{M}$ , is simply equal to  $\mu_X \lambda_A$ , implying that communities with larger population abundances are expected to be less stable and experience larger fluctuations. Interestingly, in mutualistic systems local stability implies global stability [27] and the feasibility conditions have been extensively studied [21, 23], also in connection with empirical data [28].

#### IV. ANALYTICAL SOLUTION IN THE CASE $\rho = 0$

In section III, we showed that the spectrum of  $\mathbf{M}$  is characterized by a bulk of eigenvalues and an outlier, which is determined by the mean interaction  $\mu$ . In the following, we focus on the bulk of eigenvalues, so we assume  $\mu = 0$ .

Using the cavity method [7, 29, 30], we derive in the Supplementary Information a system of equations for the spectral density of the matrix  $\mathbf{M}$ . These equations cannot be explicitly solved in the most general case, but they take a particularly simple form when the correlation  $\rho = 0$ . In this case, it is possible to write an implicit equation for the support of the spectrum, which takes the form

$$\int dx \, ds \, P_{XD}(x, s) \frac{Sx^2\sigma^2}{|\lambda - sx|^2} = 1 , \quad (12)$$

where  $P_{XD}(x, s)$  is the joint distribution of the population abundances  $x$ , with mean  $\mu_X$  and variance  $\sigma_X^2$ , and the self-regulation terms (i.e., the diagonal elements of the interaction matrix) with mean  $\mu_d$  and variance  $\sigma_d^2$ . The complex solutions  $\lambda$  of this equation define the support of the spectrum in the complex plane. In the Supplementary



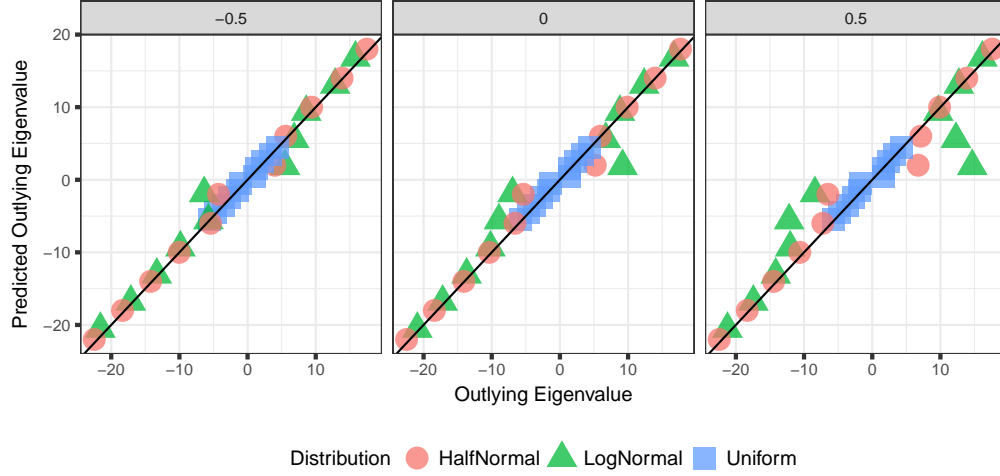


FIG. 3. The three panels show that our analytical prediction (equation 3) correctly matches the outlier of the spectrum, for different values of  $\rho$  and population abundance distributions. The matrix  $\mathbf{A}$  is built independently sampling the coefficients from a normal bivariate distribution with identical marginals defined by  $\mu$ ,  $\sigma$ , and  $\rho$ . Here, we set  $S = 1000$  and  $\sigma = 1/\sqrt{S}$ , and vary  $\mu$  between  $-10$  and  $10$  to test our prediction. We draw  $\mathbf{X}$  from three different distributions with positive support: uniform (on  $[0,1]$ ), log-normal (with mean log-mean 0.5 and log-standard deviation 0.5) and half-normal (shifted rightwards to have support  $(1, \infty)$ , and with parameter  $\theta = 1$ ).

Information we explicitly solve the case of constant self-regulation terms (i.e.,  $\sigma_d = 0$ ) and population abundances drawn from a uniform distribution.

When the self-regulation terms are constant, equation 12 reduces to

$$\int dx P_X(x) \frac{Sx^2\sigma^2}{|\lambda - \mu_d x|^2} = 1, \quad (13)$$

where  $P_X(x)$  is the species abundance distribution. Figure 4 compares the analytical prediction with the bulk of eigenvalues of  $\mathbf{M}$  for different distributions of  $\mathbf{X}$ , showing that the solutions of equation 13 closely match the support of the spectrum of  $\mathbf{M}$ . Interestingly, the integral in equation 13 is finite for any normalized  $P_X(x)$ , implying that our result holds even for very broad species abundance distribution, without any requirement on its moments (see Supplementary Information).

Equation 13 also predicts that if  $\mathbf{A}$  is stable, then  $\mathbf{M}$  is stable. In fact, equation 13 predicts that the matrix  $\mathbf{A}$  is stable iff  $\mu_d + S\sigma^2 < 0$ . If this condition is met, it is simple to observe that

$$\frac{Sx^2\sigma^2}{|\lambda - \mu_d x|^2} < 1 \quad (14)$$

for any complex  $\lambda$  with positive real part and any positive real  $x$ . When this inequality is used in equation 13, one obtains that the points on the boundary of the support, and therefore all the eigenvalues, always have negative real part.

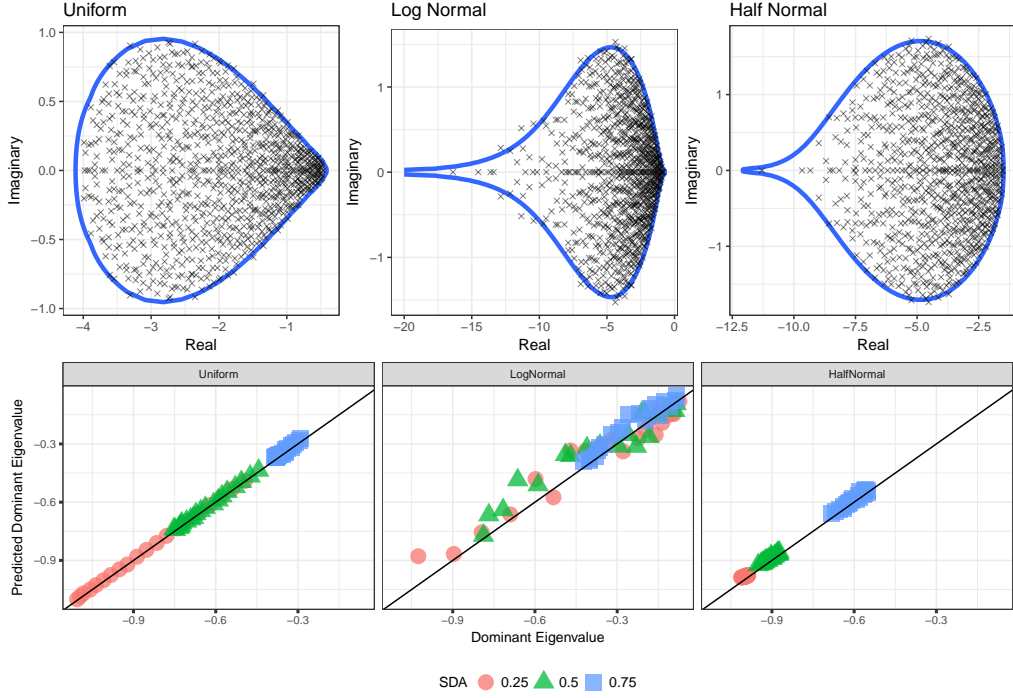


FIG. 4. The top row shows that the analytical prediction for the support of the eigenvalue distribution obtained in equation 13 (solid blue line) correctly identifies the support of the spectrum of  $M = XA$ . In all the three plots,  $A$  is built using a bivariate normal distribution with identical marginals  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$  and correlation  $\rho = 0$ . The diagonal entries of  $A$  are fixed at  $-2$ . We considered three different abundance distributions: uniform ( $X$  is sampled from a uniform distribution on  $[0.25, 1.75]$ ) lognormal ( $X$  is sampled from a log-normal distribution with log-mean 0.5 and log-standard deviation 0.5) and half-normal ( $X$  is sampled from a half-normal, shifted rightwards to have support  $(1, \infty)$ , and with parameter  $\theta = 1$ ). The bottom row shows the value of the rightmost eigenvalue of  $M$  against the analytical prediction for the leading eigenvalue of matrices with the same abundance distributions used above, but varying their variances  $\sigma_X^2$ . Different colors correspond to different values of  $\sigma$ . Each point is an average over 20 simulations. The deviations from the prediction are due to finite-size effects, which depend on the distribution of the population abundances.

## V. THE STABILITY OF LARGE COMMUNITY MATRICES DOES NOT DEPEND ON POPULATION ABUNDANCE

In the previous section, we derived the spectrum in the case  $\rho = 0$ , finding that if the interaction matrix  $A$  is stable, then  $M$  is stable. The goal of this section is to study more deeply the relationship between the stability of  $A$  and that of  $M$ . More specifically, given a stable random matrix  $A$ , we ask what is the probability of finding a positive diagonal matrix  $X$ , such that  $M = XA$  is stable.

A matrix  $A$  is *D-stable* if, for any positive diagonal matrix  $X$ ,  $XA$  is stable [31]. An explicit condition for *D-stability* that does not require checking all the possible choices of  $X$  is not known in dimension larger than four [32]. Therefore, it is not known, in general, under which values of  $\mu$ ,  $\sigma$ ,  $\rho$  and  $\mu_d$  random matrices are expected to be *D-stable*.

A stronger condition for stability is *diagonal stability*. A matrix  $A$  is diagonally stable if there exists a positive diagonal matrix  $X$  such that  $XA + A^t X$  is stable. Interestingly, diagonal stability implies *D-stability* [31]. As for

D-stability, a simple necessary and sufficient test for diagonal stability is not known. On the other hand, it is simple to observe that the stability of  $(\mathbf{A} + \mathbf{A}^t)/2$  is a sufficient condition for diagonal stability (corresponding to choosing a constant diagonal matrix  $\mathbf{X}$ ), and therefore also implies D-stability.

All the eigenvalues of  $(\mathbf{A} + \mathbf{A}^t)/2$  are real and, if  $\mathbf{A}$  is a symmetric random matrix of independently distributed entries with bounded higher moments, the bulk of eigenvalues of  $(\mathbf{A} + \mathbf{A}^t)/2$  follows Wigner's semicircle distribution [33, 34]

$$\varrho_{\frac{\mathbf{A} + \mathbf{A}^t}{2}}(\lambda) = \frac{\sqrt{2S\sigma^2(1 + \rho) - (\lambda - (\mu_d - \mu))^2}}{\pi S\sigma^2(1 + \rho)}, \quad (15)$$

with one outlying eigenvalue equal to  $\mu_d + (S - 1)\mu$ .

For positive mean  $\mu$ , if  $\mu > (1 + \rho)\sigma/\sqrt{S}$ , the rightmost eigenvalue is the outlier. In this case, the rightmost eigenvalue of  $\mathbf{A}$  and of  $(\mathbf{A} + \mathbf{A}^t)/2$  are the same. Therefore, for non-negative  $\mu$ , stable random matrices are almost surely diagonally stable. Since diagonal stability implies D-stability, if  $\mathbf{A}$  is stable, then  $\mathbf{M} = \mathbf{X}\mathbf{A}$  is stable. This argument is in agreement with our formula for the outlier of  $\mathbf{M}$  in the case of non-vanishing mean  $\mu$ , obtained in equation 11. For positive mean  $\mu$  (e.g., in mutualistic systems), the rightmost eigenvalue of  $\mathbf{M}$  is equal to  $\mu_X \lambda_{\mathbf{A}}$ , where  $\lambda_{\mathbf{A}}$  is the rightmost eigenvalue of  $\mathbf{A}$  and  $\mu_X$  is positive by definition. The sign of the rightmost eigenvalue of  $\mathbf{M}$  is therefore the same as that of the rightmost eigenvalue of  $\mathbf{A}$ .

Since a negative  $\mu$  only produces a equal shift in the rightmost eigenvalue of  $\mathbf{A}$ ,  $(\mathbf{A} + \mathbf{A}^t)/2$  and  $\mathbf{M}$ , we can restrict our analysis to the case  $\mu = 0$ . For vanishing mean, the rightmost eigenvalue of  $(\mathbf{A} + \mathbf{A}^t)/2$  is equal to [34]

$$\lambda_{\frac{\mathbf{A} + \mathbf{A}^t}{2}} = \mu_d + \sqrt{2S\sigma^2(1 + \rho)}, \quad (16)$$

which should be compared with the rightmost eigenvalue of  $\mathbf{A}$

$$\lambda_{\mathbf{A}} = \mu_d + \sqrt{S\sigma^2(1 + \rho)}. \quad (17)$$

As shown in [23, 34],  $\lambda_{\frac{\mathbf{A} + \mathbf{A}^t}{2}} \geq \lambda_{\mathbf{A}}$  and they are equal in the case  $\rho = 1$ . Equation 16 imposes a sufficient condition on diagonal stability: if

$$\mu_d + \sqrt{2S\sigma^2(1 + \rho)} < 0, \quad (18)$$

$\mathbf{A}$  is diagonally stable and, for any choice of positive diagonal matrix  $\mathbf{X}$ ,  $\mathbf{M} = \mathbf{X}\mathbf{A}$  is stable. The non-trivial regime therefore corresponds to the values of parameters where  $\mu_d + \sqrt{2S\sigma^2(1 + \rho)} > 0$  and  $\mu_d + \sqrt{S\sigma^2(1 + \rho)} < 0$  [23].

Since an explicit condition for D-stability does not exist, we computed numerically the probability that, given a stable random matrix  $\mathbf{A}$ , a positive diagonal matrix  $\mathbf{X}$  would make  $\mathbf{M} = \mathbf{X}\mathbf{A}$  unstable. Note that, since any matrix has a non-null probability of being generated when entries are sampled from a bivariate distribution with infinite support, this probability is always non-zero. The relevant question in this context is therefore how this probability depends on the number of species  $S$ . Figure 5 shows that the probability of finding a  $\mathbf{X}$  with a destabilizing effect decreases exponentially with the number of species  $S$ , with a rate that depends on the rightmost eigenvalue  $\lambda_{\mathbf{A}}$  and the correlation  $\rho$ . This implies that, for large values of  $S$ , the probability of finding an unstable matrix  $\mathbf{M}$  when  $\mathbf{A}$  is stable vanishes, and therefore, for  $S \rightarrow \infty$ ,  $\mathbf{M}$  is almost surely stable if  $\mathbf{A}$  is stable.

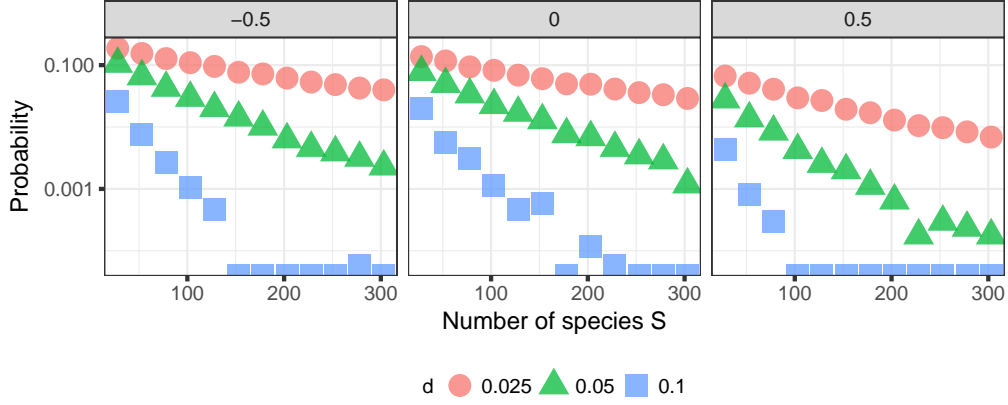


FIG. 5. We computed the probability that a matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$  is unstable (i.e., that the leading eigenvalue has positive real part), given that  $\mathbf{A}$  is a stable random matrix with rightmost eigenvalue equal to  $\lambda_{max} = -d$ . This probability decreases exponentially with  $S$  for different values of  $\rho$  ( $-0.5$  in the left panel,  $0$  in the center and  $0.5$  in the right) and  $\lambda_{max}$  (different colors). For a given number of species  $S$ , we construct the random matrix  $\mathbf{A}$  by sampling its entries from a bivariate normal with identical marginals  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$  and given  $\rho$ . The diagonal elements of  $\mathbf{A}$  are all equal and their value is determined in order to have dominant eigenvalue equal to  $\lambda_{max}$ . The diagonal entries of  $\mathbf{X}$  were sampled from a uniform distribution on  $[0, 1]$ . For each value of the parameters  $\rho$ ,  $\lambda_{max}$  and  $S$ , we constructed 15000 matrices  $\mathbf{A}$  and  $\mathbf{X}$  and computed the fraction of matrices  $\mathbf{M} = \mathbf{X}\mathbf{A}$  with positive rightmost eigenvalue.

## VI. FIXED POINTS ARE ALMOST SURELY STABLE IN LARGE RANDOM LOTKA-VOLTERRA EQUATIONS.

If we consider the Lotka-Volterra equations (equation A1), and we set the values of the intrinsic growth rates  $\mathbf{r}$ , the fixed point has components

$$x_i^* = \sum_j A_{ij}^{-1} r_j. \quad (19)$$

Let us also assume that all these components are positive (i.e.,  $\mathbf{r}$  is inside the feasibility domain). In section V, we showed that the matrix obtained by multiplying a stable random matrix  $\mathbf{A}$  and a random positive diagonal matrix  $\mathbf{X}$  is more and more likely to be stable as  $S$  increases. It is evident (from equation 19) that the components of  $\mathbf{x}^*$  are not independent of the entries of the matrix  $\mathbf{A}$ . The presence of this correlation implies that, at least in principle, choosing a random vector  $\mathbf{r}$  inside the feasibility domain to define  $\mathbf{X}$  could produce different results from sampling independent entries from a specified species abundance distribution.

In this section we repeat the simulations detailed in section V, but instead of considering a random fixed point  $\mathbf{x}^*$ , we find the  $\mathbf{x}^*$  determined by a random intrinsic growth rate vector  $\mathbf{r}$  sampled uniformly from the feasibility domain. The most intuitive method for this simulation would consist of taking a random matrix  $\mathbf{A}$ , choosing a value  $\mathbf{r}$  at random on the unit sphere, checking if it corresponds to a feasible fixed-point using equation 19, and finally computing the eigenvalue of  $\mathbf{M} = \mathbf{X}\mathbf{A}$ . However, as the number of species  $S$  increases this method becomes practically unfeasible. In fact, the fraction of intrinsic growth rate vectors  $\mathbf{r}$  corresponding to a feasible solution decreases exponentially with

$S$  [23]. If this intuitive method was employed, most of the simulation time would be spent trying to find vectors  $\mathbf{r}$  inside the feasibility domain.

On the other hand, since the relation between  $\mathbf{r}$  and  $\mathbf{x}^*$  (via equation 19) is bijective, we can easily construct all the vectors  $\mathbf{r}$  inside the feasibility domain by considering all the possible feasible solution  $\mathbf{x}^*$ . In section V we specified a distribution on the  $\mathbf{x}^*$ . This distribution translates to a non-trivial distribution on the  $\mathbf{r}$  (that can be obtained from equation 19). In this section, we instead assume a distribution on the  $\mathbf{r}$  and derive a corresponding distribution for the  $\mathbf{x}^*$ . For instance, if we assume that the vectors  $\mathbf{r}$  are uniformly distributed on the unit sphere, the distribution of the (feasible)  $\mathbf{x}^*$  reads [23]

$$P(\mathbf{x}^*|\mathbf{A}) = \frac{1}{\Xi(\mathbf{A})} \frac{\delta(\|\mathbf{x}^*\|^2 - 1)}{\|\mathbf{A}\mathbf{x}^*\|^S} \prod_{i=1}^S \Theta(x_i), \quad (20)$$

where  $\Xi(\mathbf{A})$  is a normalization factor.

Sampling vectors  $\mathbf{x}^*$  according to this distribution is equivalent to sampling vectors  $\mathbf{r}$  uniformly from the feasibility domain. It is important to observe that when  $\mathbf{x}^*$  is drawn according to this distribution, its entries are not i.i.d. independent random variables and their densities depend on  $\mathbf{A}$ .

Figure 6 shows the stability of  $\mathbf{M} = \mathbf{X}\mathbf{A}$  when the diagonal entries of  $\mathbf{X}$  are sampled from the probability distribution defined in equation 20. Despite the presence of a correlation between the entries of  $\mathbf{X}$  and  $\mathbf{A}$ , the result obtained in section V is confirmed: the probability of observing a stable  $\mathbf{A}$  but an unstable  $\mathbf{M}$  decreases exponentially with  $S$ . If the interaction matrix  $\mathbf{A}$  is stable, in the limit of large  $S$ , the set of intrinsic growth rates corresponding to feasible unstable solutions has measure zero.

In the Supplementary Information, we extended our analysis to include additional biological constraints on the intrinsic growth rates. For instance, in a predator-prey system, the intrinsic growth rates of predators are restricted to be negative, as they cannot survive in absence of their prey. As shown in the Supplementary Information, our results hold even when these biological constraints are imposed.

## VII. DISCUSSION

We explored the effect of population abundances on the stability of random interacting ecosystems. We derived an expression for the spectral density of a community matrix that explicitly includes the species abundance distribution. The effect on the eigenvalues is highly heterogeneous and strongly depends on the specific choice of the abundance distribution, with potentially important consequences on the identity of species mostly affected by the perturbation [35]. On the other hand, a remarkably simple message emerges for large randomly interacting ecosystems: the community matrix is stable if and only if the interaction matrix is stable. In other words, the abundances of species seem to not affect the sign of the eigenvalues. We further explored this intriguing result by explicitly estimating the probability of choosing a species abundance distribution leading to instability. While for finite systems this probability is always positive, it decreases exponentially with the number of species, confirming what was found studying the spectrum of the community matrix analytically.

While (local asymptotic) stability has a central role in the ecological debate, it is notoriously difficult to measure. Moreover, had we found that particular distributions of species abundance would favor stability, then we could have

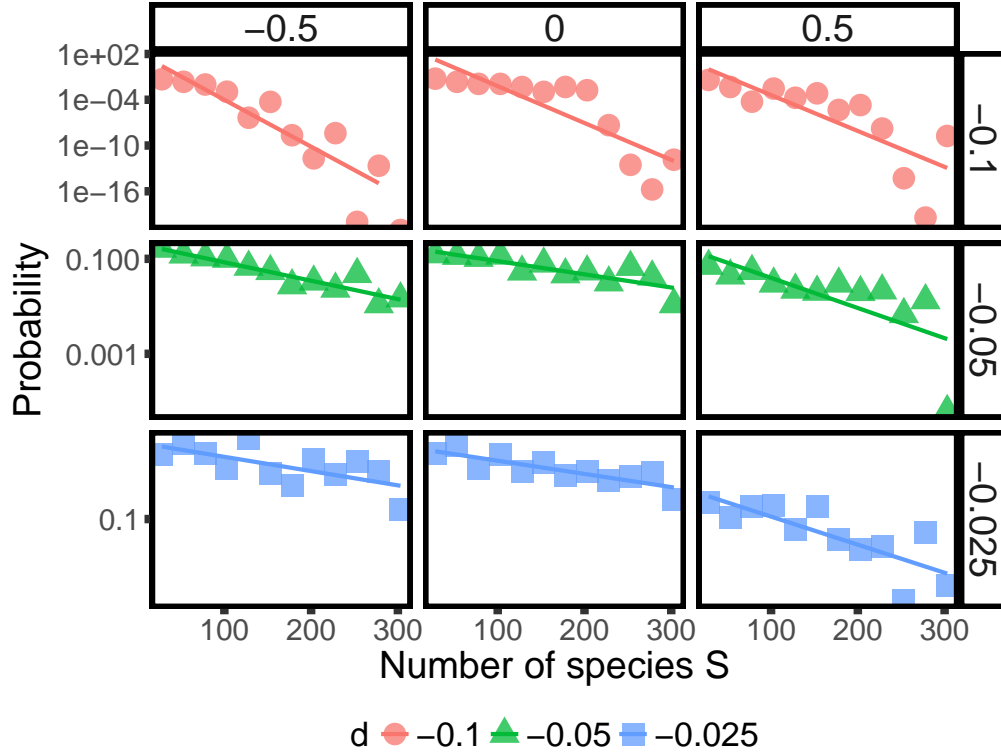


FIG. 6. These panels plot the same quantity as of fig 5. Instead of sampling  $\mathbf{X}$  from a uniform distribution, we used the distribution of eq. 20, which guarantees an unbiased sampling of the intrinsic growth rates of a Lotka-Volterra systems. This sampling method is in fact equivalent to sampling a random interaction matrix  $\mathbf{A}$  and an intrinsic growth rate vector  $\mathbf{r}$  inside the feasibility domain and checking the stability of the corresponding feasible fixed point. The exponential decay with increasing  $S$  strongly suggests that the set of feasible unstable fixed points has measure zero for large, randomly interacting Lotka-Volterra systems.

sought confirmation of our theory in the empirical literature. Given that we have found the opposite —choosing any species abundance does not affect stability—the task is obviously more complicated. Many researchers proposed that species-abundances play a key role in stabilizing communities, and maintained that the fact that May’s theory brushed them aside was one of its major drawbacks [18, 19]. Showing that this is not the case shifts the focus away from species abundances towards other mechanisms for stabilization.

Our results strongly suggest that large random matrices are D-stable *almost surely*: the set of destabilizing positive diagonal matrices has measure zero. This fact has important consequences on Lotka-Volterra systems of equations, implying that feasible unstable fixed-points are very unlikely. This result allows to disentangle the problem of feasibility (how often are fixed points feasible?) from the problem of stability (how often are fixed points stable?), justifying a-posteriori what assumed in many studies on feasibility [21, 23] and expanding the validity of their results.

A stronger notion than D-stability is diagonal stability. While for Lotka-Volterra systems, the former implies local asymptotic stability of any feasible solution, the latter implies global stability. We showed that large random stable matrices are always D-stable. Under which conditions they are also diagonally stable is an important open problem. A sufficient condition for diagonal stability is negative definiteness [23]. In the context of random matrices, negative

definiteness is equivalent to the condition expressed in equation 18. The condition for negative definiteness should be compared to the condition for stability (see equation 17). For large random matrices, two extreme scenarios are possible: negative definiteness is almost surely a necessary condition for diagonal stability, or stable random matrices are almost surely diagonally stable. It is also possible that the condition for diagonal stability is less trivial, corresponding to values of parameters between the conditions imposed by equations 18 and 17. Even more complicated, it is also possible that a sharp condition for diagonal stability does not exist for random matrices and, in the limit of large  $S$ , stable and non negative definite random matrices have a non-vanishing probability of being (or not being) diagonally stable.

Our results shed light on one of the most controversial aspects of the classic result of May [3] and its extensions. Many authors [5, 18–20, 36, 37] have argued that the unrealistic assumption of constant population abundances was a key choice in May’s paper, suggesting that more realistic abundance distribution would have produced drastically different results. We showed that the conditions obtained in the original paper and in its extension [3, 4] are in fact valid for any species abundance distribution. In other words, the stability of fixed points (i.e., the stability of the community matrix) is determined only by the stability of the interaction matrix.

## ACKNOWLEDGMENTS

We thank A. Maritan, S. Tang and G. Barabás for comments and discussions. T.G. and S.A. were supported by NSF grant DEB-1148867. J.G. was supported by the Human Frontier Science Program. T.R. is supported by the Royal Society.

- 
- [1] Arditi, R. and Ginzburg, L. R. *How Species Interact: Altering the Standard View on Trophic Ecology*. Oxford University Press, (2012).
  - [2] May, R. M. *Science* **241**(4872) (1988).
  - [3] May, R. M. *Nature* **238**(5364), 413–414 aug (1972).
  - [4] Allesina, S. and Tang, S. *Nature* **483**(7388), 205–208 feb (2012).
  - [5] Allesina, S. and Tang, S. *Population Ecology* **57**(1), 63–75 jan (2015).
  - [6] Allesina, S., Grilli, J., Barabás, G., Tang, S., Aljadeff, J., and Maritan, A. *Nature communications* **6**, 7842 jan (2015).
  - [7] Grilli, J., Rogers, T., and Allesina, S. *Nature Communications* **7** (2016).
  - [8] Bell, G. *Science* **293**(5539) (2001).
  - [9] Fisher, R., Corbet, A. S., and Williams, C. B. *The Journal of Animal Ecology* **12**(1), 42 may (1943).
  - [10] Preston, F. W. *Ecology* **29**(3), 254–283 (1948).
  - [11] Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. *Nature* **450**(7166), 45–9 nov (2007).
  - [12] Caswell, H. *Ecological Monographs* **46**(3), 327–354 feb (1976).
  - [13] Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, (2001).
  - [14] Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. *Nature* **424**(6952), 1035–7 aug (2003).
  - [15] Azaele, S., Suweis, S., Grilli, J., Volkov, I., Banavar, J. R., and Maritan, A. *Reviews of Modern Physics* **88**(3), 035003 jul (2016).
  - [16] MacArthur, R. H. *Proceedings of the National Academy of Sciences* **43**(3), 293–295 mar (1957).

- [17] Vandermeer, J. H. and MacArthur, R. H. *Ecology* **47**(1), 139–140 (1966).
- [18] James, A., Plank, M. J., Rossberg, A. G., Beecham, J., Emmerson, M., and Pitchford, J. W. *The American Naturalist* **185**(5), 680–692 may (2015).
- [19] Jacquet, C., Moritz, C., Morissette, L., Legagneux, P., Massol, F., Archambault, P., and Gravel, D. *Nature Communications* **7**, 12573 aug (2016).
- [20] Roberts, A. *Nature* **251**(5476), 607–608 oct (1974).
- [21] Rohr, R. P., Saavedra, S., and Bascompte, J. *Science* **345**(6195), 1253497–1253497 jul (2014).
- [22] Stone, L. *Nature Communications* **7**, 12857 sep (2016).
- [23] Grilli, J., Adorasio, M., Suweis, S., Barabás, G., Banavar, J. R., Allesina, S., and Maritan, A. *Nature Communications* **8**, 0 feb (2017).
- [24] Barabás, G., Michalska-Smith, M. J., and Allesina, S. *Nature Ecology & Evolution* **1**(12), 1870–1875 dec (2017).
- [25] Nguyen, H. H., O’Rourke, S., and O’Rourke, S. *International Mathematics Research Notices* **2015**(17), 7620–7689 (2015).
- [26] O’Rourke, S. and Renfrew, D. *Electronic Journal of Probability* **19** (2014).
- [27] Goh, B. S. (1979).
- [28] Fort, H. and Mungan, M. *Proceedings. Biological sciences* **282**(1808), 20150592 jun (2015).
- [29] Rogers, T., Castillo, I. P., Kühn, R., and Takeda, K. *Physical Review E* **78**(3), 031116 sep (2008).
- [30] Rogers, T. and Castillo, I. P. *Physical Review E* **79**(1), 012101 jan (2009).
- [31] Kaszkurewicz, E. and Bhaya, A. *Matrix Diagonal Stability in Systems and Computation*. Birkhäuser Boston, Boston, MA, (2000).
- [32] Redheffer, R. *SIAM Journal on Algebraic Discrete Methods* **6**(4), 612–623 oct (1985).
- [33] Wigner, E. P. *The Annals of Mathematics* **67**(2), 325 mar (1958).
- [34] Tang, S. and Allesina, S. *Frontiers in Ecology and Evolution* **2**, 21 jun (2014).
- [35] Suweis, S., Grilli, J., Banavar, J. R., Allesina, S., and Maritan, A. *Nature communications* **6**, 10179 jan (2015).
- [36] Pimm, S. L. *Oikos* **33**(3), 351 (1979).
- [37] King, A. W. and Pimm, S. L. *The American Naturalist* **122**(2), 229–239 aug (1983).
- [38] Dougoud, M., Vinckenbosch, L., Rohr, R. P., Bersier, L.-F., and Mazza, C. *PLOS Computational Biology* **14**(2), e1005988 feb (2018).
- [39] Note that this argument contradicts the proof shown in [38] (sec S4.1). We contacted the authors of the paper who agreed that a typo was present in their proof.
- [40] Rogers, T. *Journal of Mathematical Physics* **51**(9), 093304 sep (2010).



## Appendix A: Premise: stability and diagonal-stability

In the Generalized Lotka-Volterra (GLV) model:

$$\frac{dx_i(t)}{dt} = x_i(t) \left( r_i + \sum_j A_{ij} x_j(t) \right), \quad (\text{A1})$$

a feasible fixed-point (if exists) is given by

$$0 = r_i + \sum_j A_{ij} x_j^*. \quad (\text{A2})$$

The community matrix (i.e., the Jacobian evaluated at the fixed point) has elements  $M_{ij} = A_{ij} x_i^*$ .

If all the eigenvalues of  $\mathbf{M}$  have real part (i.e.,  $\mathbf{M}$  is stable), then the fixed point is stable. In general, the stability of  $\mathbf{A}$  does not guarantee that also  $\mathbf{M}$  is stable.  $\mathbf{A}$  is said to be D-stable if, for any positive diagonal matrix  $\mathbf{D}$ ,  $\mathbf{DA}$  is also stable. Clearly, by definition, D-stability implies local stability of any feasible fixedpoint. A stronger condition than D-stability is diagonal-stability.  $\mathbf{A}$  is diagonally stable if it exists a positive diagonal matrix  $\mathbf{D}$  such that  $\mathbf{AD} + \mathbf{DA}^t$  is stable. In the GLV model, if  $\mathbf{A}$  diagonally stable then the feasible fixed point (if it exists) is globally stable.

If  $\mathbf{A}$  is a random matrix, it is stable if

$$\mu_d + \max\{\sqrt{S\sigma^2}(1 + \rho), \mu S\} < 0 \quad (\text{A3})$$

where the diagonal entries of  $\mathbf{A}$  are equal to  $\mu_d$ . Each pair of off-diagonal entries  $(A_{ij}, A_{ji})$  is drawn from a bivariate distribution with identical marginal means  $\mu$ , variances  $\sigma^2$  and correlation  $\rho$ . A simple, sufficient condition for diagonal-stability is that  $a(\mathbf{A} + \mathbf{A}^t)$  is stable, which corresponds to choose a constant diagonal matrix  $\mathbf{D} = a\mathbf{I}$  in the definition of diagonal stability. If  $\mathbf{A}$  is a random matrix with parameters  $\mu, \sigma, \rho$  and  $\mu_d$ ,  $a(\mathbf{A} + \mathbf{A}^t)$  is a symmetric random matrix, with diagonal  $2a\mu_d$ , mean  $2a\mu$ , variance  $2a^2\sigma^2(1 + \rho)$  and correlation  $\rho_H = 1$ . The condition for stability for  $a(\mathbf{A} + \mathbf{A}^t)$  reads then

$$2a\mu_d + \max\{\sqrt{S2a^2\sigma^2(1 + \rho)}(1 + \rho_H), 2a\mu S\} < 0 \quad (\text{A4})$$

which simplifies into

$$\mu_d + \max\{\sqrt{2S\sigma^2(1 + \rho)}, \mu S\} < 0 \quad (\text{A5})$$

In the case when the maximum is dominated by the term with the variance, by comparing this equation with equation A3, we can easily see that if

$$(1 + \rho)^2 < \frac{\mu_d^2}{S\sigma^2} < 2(1 + \rho) \quad (\text{A6})$$

the matrix  $\mathbf{A}$ , while  $a(\mathbf{A} + \mathbf{A}^t)$  is not [39]. In this regime diagonal-stability is not trivially guaranteed and it becomes interesting to ask whether fixed points are locally stable.

## Appendix B: Notation and goals

We aim to study the spectral density of a matrix  $\mathbf{M}$  of the form  $\mathbf{M} = \mathbf{XA}$ , where  $\mathbf{X}$  is a positive diagonal matrix and  $\mathbf{A}$  a random matrix with arbitrary distribution. The diagonal entries of  $\mathbf{X}$  are drawn from an arbitrary

distribution with positive support, mean  $\mu_X$  and variance  $\sigma_X^2$ . The diagonal entries of  $\mathbf{A}$  are drawn from an arbitrary distribution with negative support, mean  $\mu_d$  and variance  $\sigma_d^2$ . Each pair of off-diagonal entries  $(A_{ij}, A_{ji})$  is drawn from a bivariate distribution with identical marginal means  $\mu$ , variances  $\sigma^2$  and correlation  $\rho$ .

Let  $\mathbf{B}$  be an  $S \times S$  random matrix with complex eigenvalues  $\lambda_i$  for  $i = 1, \dots, S$ . Its spectral density is defined as

$$\varrho(x, y) = \frac{1}{S} \sum_{i=1}^S \delta(x - \Re(\lambda_i)) \delta(y - \Im(\lambda_i)) \quad (\text{B1})$$

which, in the limit of large  $S$ , converges to

$$\varrho(x, y) = \mathbb{E}[\delta(x - \Re(\lambda_i)) \delta(y - \Im(\lambda_i))] , \quad (\text{B2})$$

where  $\mathbb{E}[\cdot]$  stands for the expectation over matrices in the ensemble.

We introduce the resolvent [40]

$$\mathcal{G}(\mathbf{q}; \mathbf{B}) = \frac{1}{S} \sum_{i=1}^S (\lambda_i - \mathbf{q})^{-1} = \frac{1}{S} \text{tr}(\mathbf{B} - \mathbf{q}\mathbf{I})^{-1} . \quad (\text{B3})$$

The variable  $\mathbf{q} = \lambda + \epsilon j$  is a quaternion (see section C for definitions and notation) and the resolvent is a function  $\mathcal{G} : \mathbb{H} \rightarrow \mathbb{H}$ .

The resolvent and the spectral density are related by the following formulas [40]:

$$\mathcal{G}(\mathbf{q}; \mathbf{B}) = \int dx dy \varrho(x, y) (x + iy - \mathbf{q})^{-1} \quad (\text{B4})$$

and

$$\varrho(x, y) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \Re \left( \frac{\partial}{\partial \bar{\lambda}} \mathcal{G}(\lambda + \epsilon j; \mathbf{B}) \right) \Big|_{\lambda=x+iy} , \quad (\text{B5})$$

where  $\frac{\partial}{\partial \bar{\lambda}}$  is the Wirtinger derivative

$$\frac{\partial}{\partial \bar{\lambda}} := \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) . \quad (\text{B6})$$

### Appendix C: Quaternions: brief review and notation

In this section we briefly review the algebra of quaternions, which is needed to efficiently perform operation with the resolvent defined in equation B3 (see [40]).

When constructing the complex numbers from the real numbers, one defines a variable  $i$  to be a root of the equation  $x^2 + 1 = 0$ . The algebraic structure of  $\mathbb{C}$  descends from the equation  $i^2 = -1$  and the algebraic structure of  $\mathbb{R}$ . Similarly, the algebra of quaternions  $\mathbb{H}$  can be defined by introducing the symbols  $i, j$  and  $k$  and the relations

$$i^2 = j^2 = k^2 = ijk = -1 . \quad (\text{C1})$$

From these equations all the multiplication rules can be obtained. In particular, it follows that multiplication in  $\mathbb{H}$  is not commutative (e.g.  $ij = -ji$ ).

A quaternion  $\mathbf{q}$  can be written as

$$\mathbf{q} = a + bi + cj + dk , \quad (\text{C2})$$

where  $a, b, c, d \in \mathbb{R}$ . Equivalently, by introducing the two complex numbers  $z = a + bi$  and  $w = c + di$  and using  $k = ij$ , one can write

$$\mathbf{q} = z + wj . \quad (\text{C3})$$

Another equivalent way to represent quaternions is to write them in matrix form

$$\mathbf{q} = \begin{pmatrix} z & w \\ \bar{w} & \bar{z} \end{pmatrix} . \quad (\text{C4})$$

It can be shown that, when written in this form, the multiplication rules of quaternions match the rules of matrix multiplication. In particular, one has that

$$(z + wj)(u + vj) = (zu - w\bar{v}) + (zv + w\bar{u})j . \quad (\text{C5})$$

We also introduce the operation

$$(z + wj) \circ (u + vj) = zu - wvj , \quad (\text{C6})$$

which, in matrix notation, corresponds to element-by-element multiplication.

The conjugate of a quaternion  $\mathbf{q} = z + wj$  is defined as  $\bar{\mathbf{q}} = \bar{z} - wj$ . From this definition, one obtains the norm of a quaternion

$$|\mathbf{q}|^2 \equiv \mathbf{q}\bar{\mathbf{q}} = |z|^2 + |w|^2 , \quad (\text{C7})$$

and the inverse

$$\mathbf{q}^{-1} \equiv \bar{\mathbf{q}} \frac{1}{|\mathbf{q}|^2} . \quad (\text{C8})$$

Moreover, the real part of a quaternion is defined as

$$\Re(\mathbf{q}) \equiv \bar{\mathbf{q}} + \mathbf{q} = \Re(z) = a . \quad (\text{C9})$$

#### Appendix D: Bulk and outliers of the spectrum of $\mathbf{M}$

In this section, we want to show that the mean of  $\mathbf{A}$  does not affect the bulk of eigenvalues of  $\mathbf{M}$ . We decompose the matrix  $\mathbf{M}$  as

$$\mathbf{M} = \mathbf{X}(\mathbf{D} - \mu\mathbf{I} + \mathbf{B} + \mu\mathbf{1}) \quad (\text{D1})$$

where  $\mathbf{I}$  is the identity matrix,  $\mathbf{1}$  is a matrix of ones and  $\mathbf{D}$  is the diagonal matrix consisting of the diagonal entries of  $\mathbf{A}$ . Written in this way,  $\mathbf{B}$  is a random matrix with mean zero and null diagonal. We will show that the bulk of the spectrum of  $\mathbf{M}$  is equivalent to the bulk of eigenvalues of the matrix  $\mathbf{J} = \mathbf{X}(\mathbf{D} - \mu\mathbf{I} + \mathbf{B})$ .

Using equation D1, the resolvent of  $\mathbf{M}$  can be written

$$\mathcal{G}(\mathbf{q}; \mathbf{M}) = \mathbb{E}\left[\frac{1}{S} \text{tr}(\mathbf{q}\mathbf{I} - \mathbf{X}\mathbf{D} - \mu\mathbf{X} - \mathbf{X}\mathbf{B} - \mu\mathbf{X}\mathbf{1})^{-1}\right] = \mathbb{E}\left[\frac{1}{S} \text{tr}(\mathbf{q}\mathbf{I} - \mathbf{J} - \mu\mathbf{X}\mathbf{1})^{-1}\right]. \quad (\text{D2})$$

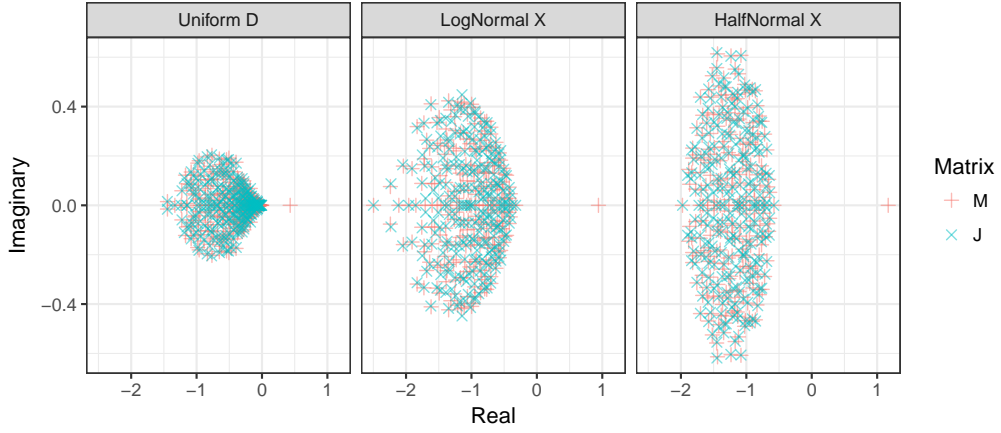


FIG. A1. The eigenvalue distributions for  $\mathbf{M}$  and  $\mathbf{J}$  of size  $S = 250$  with  $\mathbf{D}$  and  $\mathbf{X}$  following different distributions. In each case,  $\mathbf{A}$  is a bivariate normal distribution with identical marginals  $\mu = 2/S$ ,  $\sigma = 0.5$  and  $\rho = 0$ . Uniform D has  $\mathbf{D}$  following a uniform distribution on  $(-1.5, -0.5)$ . LogNormal X has  $\mathbf{X}$  following a log-normal distribution with log-mean 0 and log-standard deviation 0.35. HalfNormal X has  $\mathbf{X}$  following a half-normal distribution with support  $(1, \infty)$  and parameter  $\theta = 1$ .

Using the Sherman-Morrison formula, if  $\mathbf{Y}$  and  $\mathbf{Y} + \mathbf{Z}$  are invertible matrices and  $\mathbf{Z}$  has rank 1, then

$$(\mathbf{Y} + \mathbf{Z})^{-1} = \mathbf{Y}^{-1} + \frac{1}{1 + \text{tr}(\mathbf{Z}\mathbf{Y}^{-1})} \mathbf{Y}^{-1} \mathbf{Z} \mathbf{Y}^{-1}. \quad (\text{D3})$$

Since  $\mu \mathbf{X} \mathbf{1}$  has rank one, we have

$$(\mathbf{qI} - \mathbf{J} - \mu \mathbf{X} \mathbf{1})^{-1} = (\mathbf{qI} - \mathbf{J})^{-1} + \frac{1}{1 + \text{tr}(\mu \mathbf{X} \mathbf{1} (\mathbf{qI} - \mathbf{J})^{-1})} (\mathbf{qI} - \mathbf{J})^{-1} \mu \mathbf{X} \mathbf{1} (\mathbf{qI} - \mathbf{J})^{-1}. \quad (\text{D4})$$

By introducing the linear operator  $\langle \cdot \rangle$  defined by  $\langle \mathbf{C} \rangle = \frac{1}{S} \text{tr} \mathbf{C}$  for an  $S \times S$  matrix  $\mathbf{C}$ , we obtain

$$\mathcal{G}(\mathbf{q}; \mathbf{M}) = \mathcal{G}(\mathbf{q}; \mathbf{J}) + \frac{\mu}{1 + S\mu \langle \mathbf{X} \mathbf{1} (\mathbf{qI} - \mathbf{J})^{-1} \rangle} \langle (\mathbf{qI} - \mathbf{J})^{-1} \mathbf{X} \mathbf{1} (\mathbf{qI} - \mathbf{J})^{-1} \rangle. \quad (\text{D5})$$

In the limit of large  $S$ , the contribution from the second term in D5 is subleading. Therefore, the resolvent of  $\mathbf{M}$  converges to the resolvent of  $\mathbf{J}$  when  $S$  is large. In other words, as shown in Figure A1, the bulks of the eigenvalues of  $\mathbf{M}$  and  $\mathbf{J}$  are the same — up to finite-size corrections.

#### Appendix E: The case $\sigma = 0$

When  $\sigma = 0$ , we derive the spectrum of a matrix  $\mathbf{Q} = \mathbf{X}(\mathbf{D} + \mu \mathbf{1})$ . This case corresponds to setting  $\mathbf{B} = 0$  in equation D1. As shown in the main text, the spectral density of the matrix  $\mathbf{Q}$  is characterized by the presence of an outlier. In this section, we focus on the bulk of eigenvalues.

If we take  $\mathbf{J} = \mathbf{X}(\mathbf{D} - \mu \mathbf{I} + \mathbf{B})$  as before and set  $\sigma = 0$ , then  $\mathbf{B} = 0$ , so that  $\mathbf{M} = \mathbf{Q}$  and  $\mathbf{J} = \mathbf{X}(\mathbf{D} - \mu \mathbf{I})$ . The

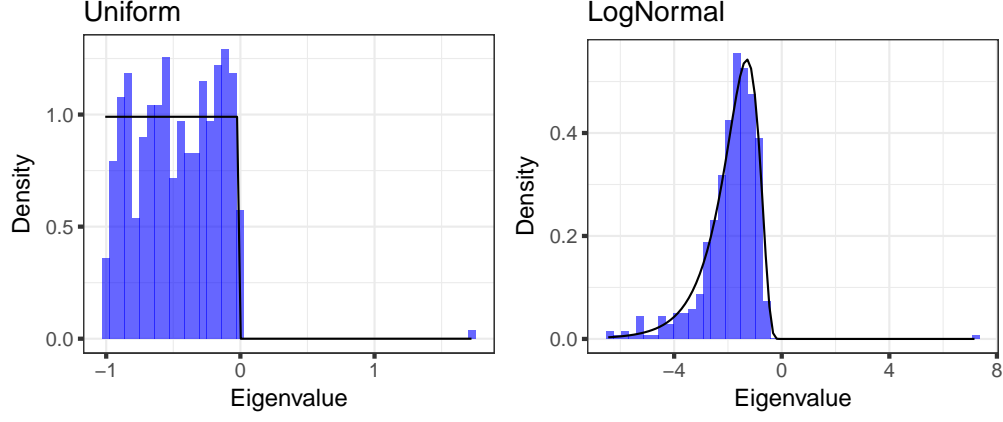


FIG. A2. Histograms of the eigenvalue distribution for two matrices  $\mathbf{Q}$  of size  $S = 1500$ . In each matrix,  $\sigma_d = 0$ ,  $\mu_d = -1$  and  $\mu = 5/S$ . The Uniform plot has  $\mathbf{X}$  following a uniform distribution on  $(0, 1)$  and the LogNormal plot has  $\mathbf{X}$  following a log-normal distribution with log-mean and log-standard deviation both 0.5.

bulk of eigenvalues of  $\mathbf{Q}$  and  $\mathbf{J}$  will be the same. The resolvent of  $\mathbf{J}$  in the case  $\sigma = 0$  reads

$$\begin{aligned}
 G(\mathbf{q}; \mathbf{J}) &= \frac{1}{S} \text{tr} (\mathbf{q} - \mathbf{J})^{-1} \\
 &= \frac{1}{S} \text{tr} (\mathbf{q} - \mathbf{X}\mathbf{D})^{-1} \\
 &= \frac{1}{S} \sum_{i=1}^S \frac{1}{q - X_i D_i}
 \end{aligned} \tag{E1}$$

since  $\mathbf{q} - \mathbf{X}(\mathbf{D} - \mu\mathbf{I})$  is symmetric. In the limit of large  $S$ , the sum in eq. E1 tends toward  $\mathbb{E} [(\mathbf{q} - \mathbf{X}\mathbf{D})^{-1}]$ . If  $P_{XD}(x, s)$  is the joint distribution of the entries of  $\mathbf{X}$  and  $\mathbf{D}$ , we obtain

$$\begin{aligned}
 G(\mathbf{q}; \mathbf{J}) &= \frac{1}{S} \sum_{i=1}^S \frac{1}{q - X_i D_i + \mu X_i} \\
 &= \mathbb{E} [(\mathbf{q} - \mathbf{X}(\mathbf{D} - \mu\mathbf{I}))^{-1}] \\
 &= \int dx ds \frac{P_{XD}(x, s)}{q - xs}.
 \end{aligned} \tag{E2}$$

In the case of a constant diagonal matrix  $\mathbf{D} = d\mathbf{I}$ , this equation simplifies to

$$G(\mathbf{q}; \mathbf{J}) = \int dx \frac{P_X(x)}{q - xd} = \frac{1}{d} \int dy \frac{P_X\left(\frac{y}{d}\right)}{q - y} \tag{E3}$$

with the change of variables  $y = xd$ . Using equation B4, we obtain that the spectral density  $\varrho_{\mathbf{J}}(\lambda)$  will be

$$\varrho_{\mathbf{J}}(\lambda) = \frac{1}{d} P_X\left(\frac{\lambda}{d}\right). \tag{E4}$$

In Figure A2, we plot the prediction from Equation E4 against the bulk of the spectrum of  $\mathbf{Q}$  for two distributions of  $\mathbf{X}$ .

## Appendix F: Derivation of the spectral density using the cavity method

In section D, we showed that we can isolate the effect of  $\mu \neq 0$ . In this section, we use the cavity method [29, 30] to derive the spectrum of the matrix  $\mathbf{J} = \mathbf{X}(\mathbf{D} - \mu \mathbf{I} + \mathbf{B})$ , where  $\mathbf{D}$  and  $\mathbf{X}$  are two random diagonal matrices and  $\mathbf{B}$  is a random matrix following the elliptic law.

We introduce the resolvent matrix

$$\mathbf{G} = (\mathbf{M} - q\mathbf{I})^{-1}, \quad (\text{F1})$$

The resolvent can be written as

$$\mathcal{G}(q; \mathbf{B}) = \frac{1}{S} \text{tr } \mathbf{G}. \quad (\text{F2})$$

Note that each element of the resolvent matrix is a quaternion. In particular we will use the notation

$$\mathbf{G}_{ik} = \alpha_{ik} + \beta_{ik}j \equiv \begin{pmatrix} \alpha_{ik} & \beta_{ik} \\ \bar{\beta}_{ik} & \bar{\alpha}_{ik} \end{pmatrix}, \quad (\text{F3})$$

while  $\mathcal{G} = \alpha + \beta j$ , where  $\alpha = \sum_i \alpha_{ii}/S$  and  $\beta = \sum_i \beta_{ii}/S$ . The cavity method allows us to compute the elements of  $\mathbf{G}$  (and therefore the resolvent  $\mathcal{G}$ ) if the matrix  $\mathbf{M}$  has a tree structure [29, 30]. It also allows to compute the spectral density for large, densely connected, random matrices [7, 29, 30]. In the limit of large  $S$ , for a densely connected matrix  $\mathbf{M}$ , the cavity equations read [7, 30]

$$\mathbf{G}_{il} \equiv \begin{pmatrix} \alpha_{il} & \beta_{il} \\ \bar{\beta}_{il} & \bar{\alpha}_{il} \end{pmatrix} = - \left( \begin{pmatrix} \lambda & \epsilon \\ \epsilon & \bar{\lambda} \end{pmatrix} + \sum_{jk} \begin{pmatrix} M_{ij} & 0 \\ 0 & M_{ji} \end{pmatrix} \begin{pmatrix} \alpha_{jk} & \beta_{jk} \\ \bar{\beta}_{jk} & \bar{\alpha}_{jk} \end{pmatrix} \begin{pmatrix} M_{kl} & 0 \\ 0 & M_{lk} \end{pmatrix} \right)^{-1}. \quad (\text{F4})$$

By introducing

$$\tilde{\mathbf{M}}_{ij} = \begin{pmatrix} M_{ij} & 0 \\ 0 & M_{ji} \end{pmatrix}, \quad (\text{F5})$$

we obtain the more compact equation

$$\mathbf{G}_{il} = - \left( \mathbf{q} + \sum_{jk} \tilde{\mathbf{M}}_{ij} \mathbf{G}_{jk} \tilde{\mathbf{M}}_{kl} \right)^{-1}. \quad (\text{F6})$$

Our goal is to find the resolvent for a random matrix of the form  $\mathbf{M} = \mathbf{X}(\mathbf{D} + \mathbf{B})$ , where  $\mathbf{X}$  and  $\mathbf{D}$  are diagonal matrices, while  $\mathbf{B}$  is a random matrix following the elliptic law. We introduce the matrix

$$\begin{pmatrix} -q\mathbf{I} & \tilde{\mathbf{X}} \\ -\tilde{\mathbf{D}} - \tilde{\mathbf{B}} & \mathbf{I} \end{pmatrix}, \quad (\text{F7})$$

which, when quaternions are represented as  $2 \times 2$  matrices, is a  $4S \times 4S$  matrix. In particular, this matrix is composed of  $S^2$   $4 \times 4$  blocks with entries

$$\begin{pmatrix} -\lambda\delta_{ij} & -\epsilon\delta_{ij} & X_{ii}\delta_{ij} & 0 \\ -\epsilon\delta_{ij} & -\bar{\lambda}\delta_{ij} & 0 & X_{ii}\delta_{ij} \\ -D_{ii}\delta_{ij} - B_{ij} & 0 & \delta_{ij} & 0 \\ 0 & -D_{ii}\delta_{ij} - B_{ji} & 0 & \delta_{ij} \end{pmatrix}. \quad (\text{F8})$$

It is simple to observe that

$$\mathbf{T} = \begin{pmatrix} -\mathbf{q} & \mathbf{X} \\ -\mathbf{A} & \mathbf{I} \end{pmatrix}^{-1} = \begin{pmatrix} (-\mathbf{q}\mathbf{I} + \mathbf{X}(\mathbf{D} + \mathbf{B}))^{-1} & \dots \\ \dots & \dots \end{pmatrix} = \begin{pmatrix} \mathbf{G} & \dots \\ \dots & \dots \end{pmatrix}. \quad (\text{F9})$$

If we write the cavity equation for  $\mathbf{T}$ , assuming dense matrices, we obtain

$$\mathbf{T}_{ii} = \left[ \begin{pmatrix} -\mathbf{q} & \tilde{\mathbf{X}}_{ii} \\ -\tilde{\mathbf{D}}_{ii} & \mathbf{I} \end{pmatrix} - \sum_{j,k} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}_{ij} & \mathbf{0} \end{pmatrix} \mathbf{T}_{jk} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}_{ki} & \mathbf{0} \end{pmatrix} \right]^{-1}. \quad (\text{F10})$$

We can apply the law of large numbers to take the expectation of the matrices over  $\mathbf{B}$ . Using the following expectations over the elements of  $\mathbf{B}$

$$\begin{aligned} \mathbb{E}[B_{ij}] &= 0 \\ \mathbb{E}[(B_{ij})^2] &= \frac{\tilde{\sigma}^2}{S} \\ \mathbb{E}[B_{ij}B_{ji}] &= \rho \frac{\tilde{\sigma}^2}{S}, \end{aligned} \quad (\text{F11})$$

we find that the non-diagonal terms of  $\mathbf{T}$  go to zero in expectation. By introducing the notation

$$\mathbf{T}_{ii} \equiv \begin{pmatrix} \mathbf{G}_{ii} & \mathbf{T}_{ii}^b \\ \dots & \dots \end{pmatrix} \quad (\text{F12})$$

and

$$\mathbf{T}_\star \equiv \frac{1}{S} \sum_i \mathbf{T}_{ii} = \begin{pmatrix} \mathcal{G} & \mathbf{T}_\star^b \\ \dots & \dots \end{pmatrix}, \quad (\text{F13})$$

the cavity equations for the diagonal terms read

$$\mathbf{T}_{ii} = \begin{pmatrix} -\mathbf{q} & X_{ii} \\ -D_{ii} - \mathbf{t} \circ \mathbf{T}_\star^b & \mathbf{I} \end{pmatrix}^{-1} \quad (\text{F14})$$

where  $\mathbf{T}_\star = \sum_i \mathbf{T}_{ii}/S$ , while  $\mathbf{t} = \tilde{\sigma}^2 \rho + \tilde{\sigma}^2 \mathbf{j}$  with  $\circ$  denoting the element-by-element matrix product (see section C).

We obtain a system of  $2S$  quaternionic equations

$$\begin{cases} \mathbf{T}_{ii}^b = -X_{ii} \left( -\mathbf{q} - X_{ii} \left( -D_{ii} - \mathbf{t} \circ \mathbf{T}_\star^b \right) \right)^{-1} \\ \mathbf{G}_{ii} = \left( -\mathbf{q} - X_{ii} \left( -D_{ii} - \mathbf{t} \circ \mathbf{T}_\star^b \right) \right)^{-1}, \end{cases} \quad (\text{F15})$$

where

$$\mathbf{T}_\star^b = \frac{1}{S} \sum_i \mathbf{T}_{ii}^b. \quad (\text{F16})$$

Assuming that the elements of  $\mathbf{X}$  and  $\mathbf{D}$  are drawn from a given joint distribution  $P_{XD}$ , we obtain the following two equations

$$\begin{cases} \mathbf{T}_\star^b = - \int dx ds P_{XD}(x, s) x \left( -\mathbf{q} + x(s + \mathbf{t} \circ \mathbf{T}_\star^b) \right)^{-1} \\ \mathcal{G} = \int dx ds P_{XD}(x, s) \left( -\mathbf{q} + x(s + \mathbf{t} \circ \mathbf{T}_\star^b) \right)^{-1}. \end{cases} \quad (\text{F17})$$

At this point one can use  $\mathbf{T}_*^b = \alpha^b + \beta^b j$  and  $\mathcal{G} = \alpha + \beta j$  to obtain 4 equations of complex variables. It is simple to observe that  $\beta^b = 0$  is always a solution and  $\beta = 0$  if and only if  $\beta^b = 0$ . The solution  $\beta = \beta^b = 0$  always corresponds to a null spectral density [40]. The values of  $\lambda$  for which a non-zero solution for  $\beta$  exist correspond to the support of the spectral density. Setting  $\beta^b \neq 0$  and  $\epsilon = 0$ , one obtains from equation F17

$$\begin{cases} \alpha^b = \int dx ds P_{XD}(x, s) \frac{x(\bar{\lambda} - sx - x\tilde{\sigma}^2 \rho \bar{\alpha}^b)}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2 + |x\tilde{\sigma}^2 \beta^b|^2} \\ 1 = \int dx ds P_{XD}(x, s) \frac{x^2 \tilde{\sigma}^2}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2 + |x\tilde{\sigma}^2 \beta^b|^2} \end{cases} \quad (\text{F18})$$

By using the second equation, the system of equations further simplifies to

$$\begin{cases} \alpha^b + \rho \bar{\alpha}^b = \int dx ds P_{XD}(x, s) \frac{x(\bar{\lambda} - sx)}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2 + |x\tilde{\sigma}^2 \beta^b|^2} \\ 1 = \int dx ds P_{XD}(x, s) \frac{x^2 \tilde{\sigma}^2}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2 + |x\tilde{\sigma}^2 \beta^b|^2} \end{cases} \quad (\text{F19})$$

The values of  $\lambda$  for which a solution of equations F19 exists are contained in the support of the spectral density. We assume that the solution  $\beta^b$  of these equations vanishes at the boundaries of the support. In this case, the points at the boundaries of the support of the spectral density are the complex solutions  $\lambda$  of

$$\begin{cases} \alpha^b + \rho \bar{\alpha}^b = \int dx ds P_{XD}(x, s) \frac{x(\bar{\lambda} - sx)}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2} \\ 1 = \int dx ds P_{XD}(x, s) \frac{x^2 \tilde{\sigma}^2}{|\lambda + x(-s - \alpha^b \rho \tilde{\sigma}^2)|^2} \end{cases} \quad (\text{F20})$$

where we used the second equation

We should note here that this method does not prove the convergence in any mode of the spectral bulk, but does yield a prediction for its support.

### SUPPORT OF THE SPECTRAL DENSITY IN THE CASE $\rho = 0$

In the case  $\rho = 0$ , the two equations F20 become independent, and the support is defined by the solutions of

$$1 = \tilde{\sigma}^2 \int dx ds P_{XD}(x, s) \frac{x^2}{|\lambda - sx|^2} \quad (\text{F21})$$

In the case  $\tilde{\sigma}_d^2 = 0$ , this equation further simplifies to

$$1 = \tilde{\sigma}^2 \int dx P_X(x) \frac{x^2}{|\lambda - \mu_d x|^2} \quad (\text{F22})$$

For instance, if  $P_X$  is a uniform distribution on  $[0, 1]$ , one can evaluate the integral, obtaining

$$\begin{aligned} 1 &= \int_0^1 dx \frac{x^2 \tilde{\sigma}^2}{|\lambda - \mu_d x|^2} \\ &= \frac{\tilde{\sigma}^2}{\mu_d^3} \left( 2\lambda \mu_d - \mu_d^2 + 2\lambda(\mu_d - \lambda) \log \left| \frac{\lambda}{\lambda - \mu_d} \right| \right) \end{aligned} \quad (\text{F23})$$

The complex solutions  $\lambda$  of this equation define the boundary of the support of the spectral density.

It is interesting to observe that the integral that defines the boundary of the support converge for any distribution  $P_X(x)$ , independently of whether the moments of  $P_X(x)$  are finite. For example, one can consider

$$P_X(x) = \frac{2}{\pi} \frac{\gamma}{(x - x_m)^2 + \gamma^2} \Theta(x - x_m) \quad (\text{F24})$$



where  $x_m$  and  $\gamma$  are two positive parameters, while the theta function  $\Theta(x - x_m)$  defines  $[x_m, \infty)$  as the support of the distribution. It is simple to observe that, since this distribution has a power-law tail with exponent  $-2$ , all the moments (including the mean) diverge. On the other hand, the integral in equation F22 converges. Figure A7 compares the analytical result obtained in equation F22 with numerical simulations for the power-law distribution of population abundances defined in equation F24.

### CONSTRAINING THE SIGN OF THE INTRINSIC GROWTH RATES

The only constraint that we have imposed on the population abundances, is that they are positive (i.e. feasible). In the context of Lotka-Volterra equations, this is equivalent to consider all the possible combinations of intrinsic growth rates inside the feasibility domain. On the other hand, other constraints than feasibility can further restrict the range of biologically-realistic growth rates. For instance, in a predator-prey system, the intrinsic growth rates of predators are constrained to be negative, since predators cannot survive in the absence of their prey.

Here we explicitly consider a predator-prey system in order to test whether these additional constraints can affect our results. We focus on a bipartite predator-prey system. In this case the interaction matrix has the following block structure:

$$\mathbf{A} = \left( \begin{array}{c|c} \mathbf{C_P} & \mathbf{B_P} \\ \hline \mathbf{B_A} & \mathbf{C_A} \end{array} \right), \quad (\text{F25})$$

where the first  $S_P$  entries of the vector of abundance correspond to prey and the last  $S_A$  entries correspond to predators. The square matrices  $\mathbf{C_P}$  ( $\mathbf{C_A}$ ) is the competition matrix between prey (predators). The  $S_P \times S_A$  matrix  $\mathbf{B_P}$  measures the rate of predation, while the  $S_A \times S_P$  matrix  $\mathbf{B_A}$  measures the growth of the predators thanks to predation. By definition, all the elements of the matrices  $\mathbf{C_P}$ ,  $\mathbf{C_A}$  and  $\mathbf{B_P}$  are negative, while the elements of  $\mathbf{B_A}$  are positive.

We considered two alternative scenarios. In one case, we considered all the feasible solutions, by sampling at random species abundances (unconstrained scenario). In the second case, we considered only feasible solutions corresponding to positive growth rates for prey and negative for predators (constrained scenario). For both the constrained and the unconstrained scenario, we computed the probability of finding a destabilizing vector of abundances (analogously to Figure 5). Figure A8 shows that the probability of finding a destabilizing vector decrease exponentially with the number of species in both the unconstrained and the constrained scenario, strongly suggesting that adding these biological constraints to the growth rates does not affect our results. It is important to notice that the exponential decay for the constrained scenario would trivially follow from the unconstrained case only if a constant fraction of feasible growth rates respected the sign constraints. As shown in Figure A8 this is not the case: an exponentially decaying fraction of growth rates vectors have biologically realistic signs.

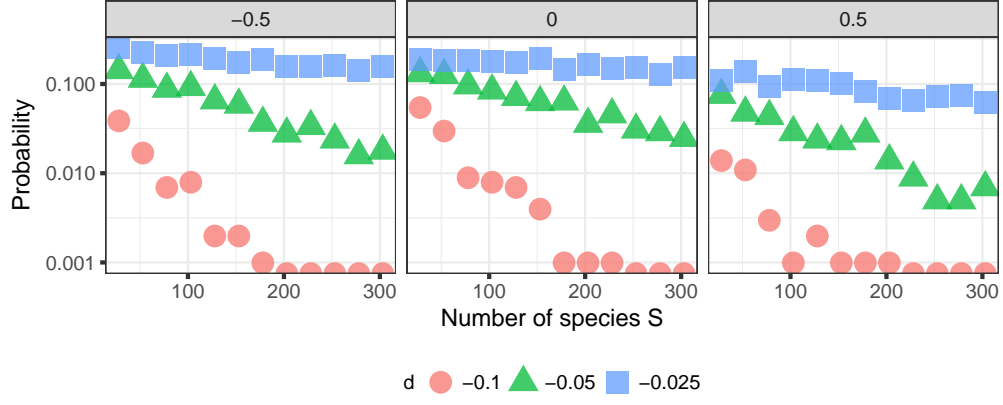


FIG. A3. The conditional probability of a matrix  $\mathbf{M}$  becoming *stable* given an *unstable* interaction matrix  $\mathbf{A}$  with leading eigenvalue  $-d$ . The off-diagonal elements of  $\mathbf{A}$  follow a bivariate normal distribution and, for each  $S$ ,  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$  and  $\rho = -0.5, 0$  or  $0.5$ . The elements of  $\mathbf{X}$  are drawn from a uniform distribution on  $(0, 1)$ , so that  $\mu_X = 0.5$  and  $\sigma_X^2 = \frac{1}{12}$ . The diagonal elements of  $\mathbf{A}$  are fixed at  $-1$  since  $\sigma_d^2 = 0$  and  $\mu_D = -1$ . Each probability is calculated from 1000 trials from 1000 simulated matrices  $\mathbf{M} = \mathbf{X}\mathbf{A}$ .

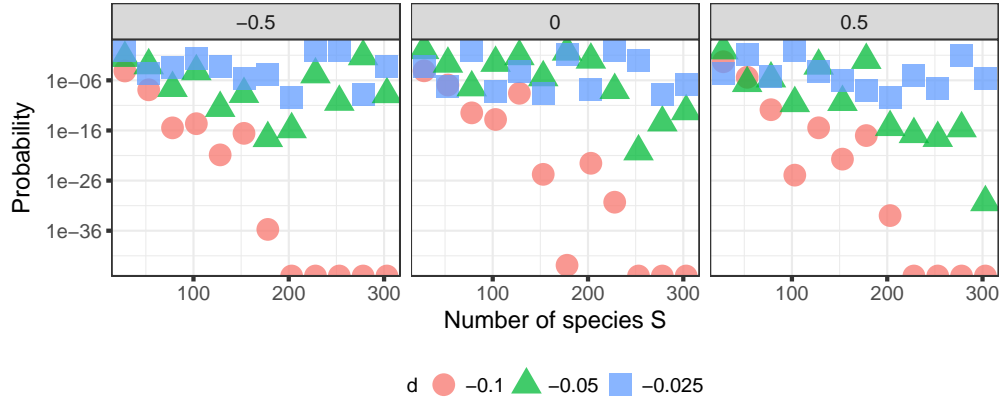


FIG. A4. As in Figure A3, except now the conditional probability is weighted by feasibility.

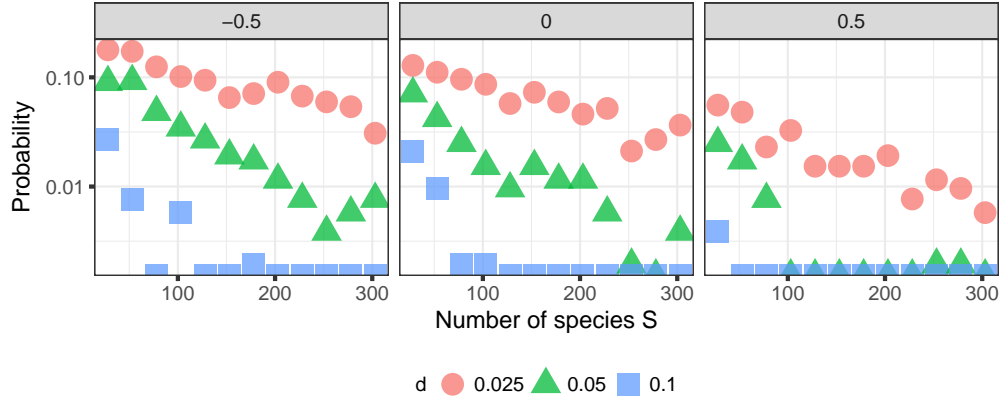


FIG. A5. The conditional probability of a matrix  $\mathbf{M}$  becoming unstable given a stable interaction matrix  $\mathbf{A}$  with leading eigenvalue  $-d$ . The off-diagonal elements of  $\mathbf{A}$  follow a bivariate normal distribution and, for each  $S$ ,  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$  and  $\rho = -0.5, 0$  or  $0.5$ . The elements of  $\mathbf{X}$  are drawn from a uniform distribution on  $(0, 1)$ , so that  $\mu_X = 0.5$  and  $\sigma_X^2 = \frac{1}{12}$ . The diagonal elements of  $\mathbf{A}$  follow a uniform distribution on  $(-0.75, -1.25)$  so that  $\sigma_d^2 = \frac{1}{48}$  and  $\mu_D = -1$ . Each probability is calculated from 1000 trials from 1000 simulated matrices  $\mathbf{M} = \mathbf{X}\mathbf{A}$ .

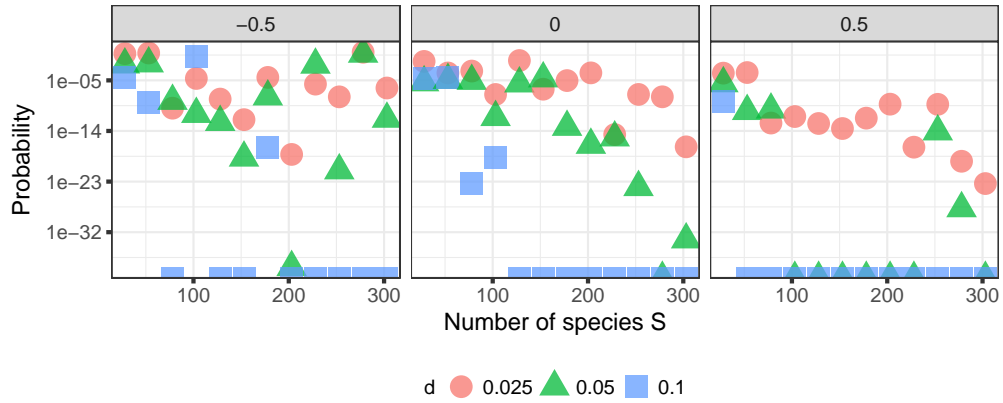


FIG. A6. As in Figure A5, except now the conditional probability is weighted by feasibility.

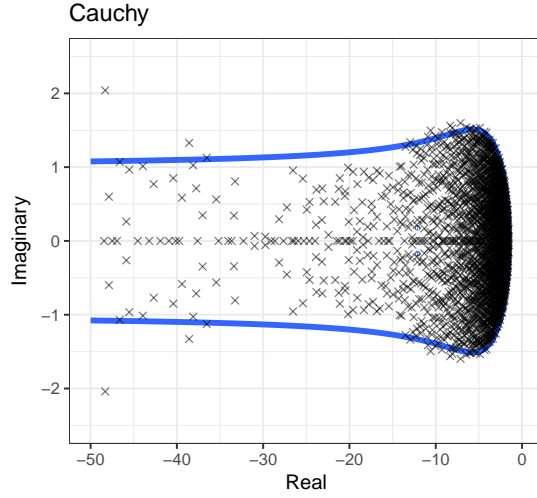


FIG. A7. The analytical predictions for the support of the eigenvalue distribution obtained in equation F22 (solid blue line) correctly predict the support of the spectrum of  $\mathbf{M} = \mathbf{X}\mathbf{A}$  for Cauchy-distributed diagonal elements of  $\mathbf{X}$ . In all the three plots,  $\mathbf{A}$  is built using a bivariate normal distribution with identical marginals  $\mu = 0$ ,  $\sigma = 1/\sqrt{S}$  and correlation  $\rho = 0$ . The diagonal entries of  $\mathbf{A}$  are fixed at  $-1$ .  $\mathbf{X}$  is sampled from the distribution defined in equation F24 with parameters  $x_m = 1$  and  $\gamma = 1$ . Our analytical calculation correctly predicts the boundaries of the spectrum even if none of the moments of  $P_X(x)$  is finite.

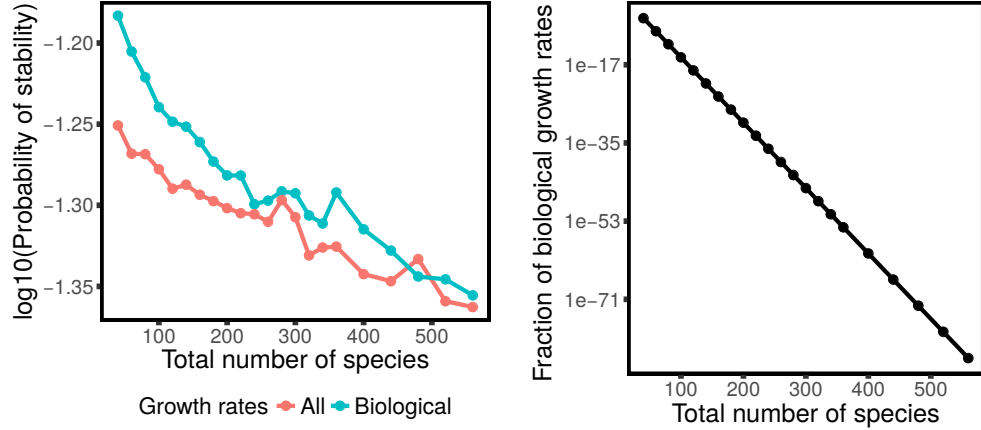


FIG. A8. The left panel shows probability that a matrix  $\mathbf{M} = \mathbf{X}\mathbf{A}$  is unstable (i.e., that the leading eigenvalue has positive real part), given that  $\mathbf{A}$  is a stable random matrix with rightmost eigenvalue equal to  $\lambda_{max} = -0.01$ . The interaction matrix  $\mathbf{A}$  is parametrized as shown in equation F25. We considered all the feasible population abundances (blue points), and only population abundances corresponding to positive growth rates for prey and negative for predators (red points). In both the case we observe an exponentially decaying fraction of destabilizing vectors. The right panels shows that the fraction of biological growth rates is also exponentially decreasing, making the result shown in the left panel non trivial. Here we considered uniformly distributed population abundances and constant diagonal competition matrices. The entries of  $\mathbf{B}_P$  were drawn from a uniform distribution with mean  $-1$  and standard deviation equal to  $0.5$ . The entries of  $\mathbf{B}_A$  were drawn from a uniform distribution with mean  $0.25$  and standard deviation equal to  $0.125$ .